



Spacetime Characterization of Real-Time Collaborative Editing

Gabriele d'Angelo, Angelo Di Iorio, Stefano Zacchioli

► To cite this version:

Gabriele d'Angelo, Angelo Di Iorio, Stefano Zacchioli. Spacetime Characterization of Real-Time Collaborative Editing. Proceedings of the ACM on Human-Computer Interaction , Association for Computing Machinery (ACM), 2018, CSCW, 2, pp.1-19. 10.1145/3274310 . hal-01882069

HAL Id: hal-01882069

<https://hal.inria.fr/hal-01882069>

Submitted on 26 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spacetime Characterization of Real-Time Collaborative Editing

GABRIELE D'ANGELO, Dept. of Computer Science and Engineering, University of Bologna, Italy
 ANGELO DI IORIO, Dept. of Computer Science and Engineering, University of Bologna, Italy
 STEFANO ZACCHIROLI, IRIF, University Paris Diderot & Inria, France

Real-Time Collaborative Editing (RTCE) is a popular way of instrumenting cooperative work on documents, in particular on the Web. Little is known in the literature yet about RTCE usage patterns in the real world. In this paper we study how a popular RTCE editor (Etherpad) is used in the wild, digging into the edit histories of a large collection of documents (about 14 000 *pads*), retrieved from one of the most popular public instances of the platform, hosted by the Wikimedia Foundation.

The pad analysis is supported by a novel conceptual model that allows to label edit operations as “collaborative” or not depending on their distance—in edit position (*space*), edit *time*, or *spacetime* (both)—from edits made by other authors. The model is applied to classify all edits from the pad corpus. Classification results are further used to characterize the collaboration behavior of pad authors.

Findings show that: 1) about half of the pads have a single author and hence witnessed no collaboration; 2) collaboration on common document parts happens often, but it happens asynchronously with authors taking turns in editing; and 3) simultaneous editing of common document parts happens very rarely. These findings help in revisiting early RTCE design decisions (e.g., the granularity of conflict management in RTCE protocols) and give insights on how to address novel needs (e.g., end-to-end encryption and offline editing).

CCS Concepts: • **Human-centered computing** → *Collaborative and social computing systems and tools*;

Additional Key Words and Phrases: collaborative editing, real-time, user behavior, conflict resolution, simultaneous editing, etherpad

ACM Reference Format:

Gabriele D'Angelo, Angelo Di Iorio, and Stefano Zacchiroli. 2018. Spacetime Characterization of Real-Time Collaborative Editing. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW, Article 41 (November 2018), 19 pages. <https://doi.org/10.1145/3274310>

1 INTRODUCTION

After being mainly a research topic for many years, Real-Time Collaborative Editing (RTCE) have also spawned tools for the masses thanks to the diffusion of platforms like Google Docs [14], Apache Wave [31], and Etherpad [12]. These platforms empower users to work *collaboratively* on the same document and *simultaneously*, with minimal setup via Software-as-a-Service deployment models. The underlying concurrency models are generally based on Operational Transformation (OT) [10], which accounts for consistency maintenance therefore relieving the users from having to worry about conflicts and their resolution.

Authors' addresses: Gabriele D'Angelo, Dept. of Computer Science and Engineering, University of Bologna, Italy, g.dangelo@unibo.it; Angelo Di Iorio, Dept. of Computer Science and Engineering, University of Bologna, Italy, angelo.diiorio@unibo.it; Stefano Zacchiroli, IRIF, University Paris Diderot & Inria, France, zack@irif.fr.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2018/11-ART41 \$15.00
<https://doi.org/10.1145/3274310>

While RTCE *can* be used by multiple authors to edit the very same parts of a document at the same time, it can also be used for non-simultaneous collaboration by a group of authors (with authors “taking turns” at editing the same parts of a document); or by a single author to share documents with multiple readers (who do not make any change on the document); or even by a single author just to work on a document from multiple devices.

Even when there actually is collaboration among multiple authors, the collaboration can happen along multiple independent dimensions. For instance, nearby parts of a document can be edited collaboratively by multiple users—we call this *space collaboration*—simultaneously or not. As opposite, collaborators can edit in a very short time frame the same document—*time collaboration*—no matter *where* their edits happen in the text. Note that only when the collaboration is both simultaneous and happening on nearby document parts—*spacetime collaboration*—there is a need of sophisticated and very fine grained concurrency models.

But what are the most typical user behavior patterns in RTCE? Specifically, to what extent there is space-, time-, and spacetime-collaboration in real-world applications of RTCE technologies?

Answers to these questions will bring a better understanding of how people edit documents collaboratively. Such an understanding will, in turn, allow to critically review early design decisions in RTCE for appropriateness with respect to actual use, as well as proceed to informed design of novel RTCE functionalities.

The goal of this work is to provide a first set of answers by doing a large-scale quantitative analysis of RTCE editing sessions in the real world. To the best of our knowledge such an analysis, at this scale, is novel in the literature.

We introduce a framework—both theoretical and implemented as a set of tools—to classify edit operations performed using the EtherPad editor as either “collaborative” or “non collaborative” along three dimensions: space, time, and spacetime. We then studied the full editing histories of about 14 000 textual documents (or *pads*, in EtherPad terminology) from <http://etherpad.wikimedia.org/>, which is one of the most popular public instances of Etherpad, hosted by the Wikimedia Foundation.¹ In particular, we answer the following research question for the given pad corpus:

To what extent edit operations happen collaboratively among authors? Specifically, to what extent do they happen:

RQ1) ...at about the same time of operations by other authors? I.e., *which edits are time-collaborative?*

RQ2) ...nearby operations by other authors? I.e., *which edits are space-collaborative?*

RQ3) ...nearby, in both space and time, operations by other authors? I.e., *which edits are spacetime-collaborative?*

Our findings show that:

- (1) about half of the pads have been edited by a single author, witnessing no collaboration whatsoever;
- (2) space-collaboration is commonplace, but is usually asynchronous, with authors taking turns in editing co-authored document parts;
- (3) spacetime-collaboration happens very rarely, calling into question the need of very fine-grained conflict resolution mechanisms, document models, and user interfaces;

Paper structure. Section 2 reviews related work and in particular previous analyses on collaborative editing. Section 3 introduces very briefly the conceptual framework used to classify edits,

¹The Wikimedia Foundation is the nonprofit organization that is home of the Wikipedia free encyclopedia and related projects. See <https://wikimediafoundation.org/>.

which is fully formalized in Appendix A. Sections 4 to 7 discuss our experiments, findings, threats to validity, and implications. A discussion of future work concludes the paper in Section 8.

2 RELATED WORK

Several experimental studies have been conducted in the past on collaborative editing. We review them here, comparing goals and methodologies with those of this paper. Previous findings will be compared with ours in Section 6.

In this paper we focus on *active editing* of document content by multiple authors, rather than collaborative writing in a more general sense. For instance authors can use highlight/comment features (e.g., as supported by Google Docs) or instant messaging to discuss what and how to write. Posner and Baecker [26] studied these collaboration patterns and identified collaborators roles such as who scribes while someone else dictates; or a peer-reviewer who provides comment to the main author via out-of-document channels. These interactions are not captured by our framework whose focus is document editing.

Specifically, we study the *proximity* of edit operations, in space and/or time. The classification of collaboration over these two dimensions is not new. The seminal work of Johansen [18] introduced a time/space matrix for characterizing collaborative actions that distinguishes between physical space shared by the authors vs remote collaboration (space) and synchronous vs asynchronous writing (time). Our notion of space is different: it refers to document areas rather than the author location in the real world. Several extensions of Johansen’s framework has been proposed to classify collaborative tools—for instance by Grudin [15] who also considered user knowledge of the places and times in which collaboration happens or by Andriessen [11] who conceptualized different collaboration processes and interactions among users. All these models look at collaboration from a more general perspective than ours, without characterizing individual edit operations.

A popular approach to study collaborative editing consists in interviewing users. Chu and Kenney [27] collected feedback from 22 undergraduates who used both MediaWiki and Google Docs for a class; all students appreciated the real-time aspect of Google Docs, but they considered MediaWiki a comparable platform in terms of collaboration enablement. Brodhal et al. [6] conducted a larger interview-based study, collecting feedback from 166 students; most students reported that RTCE editing did *not* contribute to increase the quality of collaboration, with almost 70% of them considered RTCE functionalities unsatisfactory. Participants would have preferred to be able to just comment on content, instead of being able to modify it concurrently.

The interviews-based approach is a key difference with our work, since we processed edit logs *a posteriori* without any direct involvement of the authors. The fact that these experiments involved students is a further element of contrast: we had no information about user profiles or group dynamics. Other focused on these aspects; for instance, Birnholtz and Ibara [4] studied how people “*perceive and consider the potential impacts of their own and others’ edits as they write together*” concluding that people are heavily influenced by the team organization and tend to ask themselves what their collaborators (and managers in particular) would think of their writing.

In a recent work [33], Wang et al. interviewed 30 participants from both academia and industry about the adoption of commercial RTCE tools. Participants agreed that RTCE features are valuable for collaborative writing, but they pointed out that social, personal, and privacy issues remain important barriers to the adoption of these tools. The interviews by Olenewa et al. [23] also shared insights about the advantages and inconveniences of using Google Docs. The authors identified collaboration patterns and subsequently identified issues, both organizational and technical, in properly planning editing activities among collaborators, that should be addressed to fully reap the benefits of RTCE.

Quantitative analyses of RTCE edit histories for characterizing collaboration behavior remain scarce in the literature. Birnholtz et al. [5] conducted a laboratory experiment with 150 students to study how student pairs use Google Docs, with or without the integrated chat, for both synchronous and asynchronous collaborative editing. The main goal of the study was to investigate if and to what extent people communicate while writing collaboratively. The experiment relied on both questionnaires and analysis of three types of logs: edits, chat messages, and document comments. The edits were collected by parsing Google Docs revision histories, and classifying them as major or minor, according to the amount of affected text. The authors collected data for different groups of users and measured their correlation. Their findings suggest that “*communication was useful, but had a positive relational impact primarily in synchronous writing*” and that “*text communication complemented, rather than substituted for, editing*”. Focusing on editing, the authors concluded that too much editing can negatively impact collaboration and that more communication leads to more editing activities. The research question of Birnholtz et al. work and ours differ: we focus on edit operations and their locality (in space and time), rather than studying the relationship between in-document and out-of-document interactions. The size of our corpus is significantly larger than theirs, but is limited to textual documents. Finally, our evaluation is on a set of documents fully open and not under our control, while their experiment happened in a controlled environment with pre-defined editing tasks.

Olson et al. [24] recently analyzed Google Docs logs in order to study how student teams write collaboratively. The authors collected the revision histories of 96 documents edited by some students over a 3-year period. The paper introduced the notions of *slice*, defined as a snapshot of the document, and *session*, defined as a set of slices. Since slices are associated to timestamps, it was possible to aggregate them into sessions, and to study the amount of collaborative sessions. The authors also exploited a sophisticated interface (DocuViz [32]) to visually analyze collaboration writing patterns. The experiment shows that the amount of simultaneous writing increases when authors knew and trust each other, and that the quality of produced documents is positively affected by collaborative writing.

Our research question is different from Olson et al. one, as no notion of document quality or trust between authors was directly applicable to our corpus. Also, as it was the case for Birnholtz et al., there are significant differences in corpus size, RTCE tool (Etherpad vs Google Docs), and controlled vs open platform. The methods used to measure space collaboration are comparable though. Both works use the number of characters in the final document written by each author as a measure of participation. In both cases, the deleted content is not taken into account but the contribution of each author is captured correctly.

A controlled experiment on synchronous vs asynchronous editing with Etherpad was performed by Andre et al. [3]. The authors asked a limited set of testers (up to 4 authors per group) to edit collaboratively the same document adopting two different strategies—sequential writing (some users created a first draft and other revised it) and real-time editing (all users edited the same document together)—and compared the results. The size and goals of the experiment is different from ours, even if both used Etherpad. Andre et al. focused on the time dimension and the interaction among authors, while we have also analyzed space (and spacetime) collaboration. They also studied the quality of the output, an aspect which is not taken into account by us. Finally, our study is on a much larger dataset collected from an open platform, while theirs was highly controlled.

Controlled experiments have also been used for evaluating the performances of RTCE protocols and algorithms [2, 17, 20, 25, 28]. These experiments target a lower level of abstraction of RTCE with respect to ours. They are meant to measure the efficiency of RTCE implementations in resolving conflicts and propagating data among clients, while our goal is to study how people collaborate regardless of the underlying low-level implementation.

3 CHARACTERIZING REAL-TIME COLLABORATION

In this section we briefly introduce the model we used to study real-time collaboration on Etherpad [12]. Interested readers can find full details in Appendix A.

Etherpad supports textual documents only, called *pads*. When a pad is opened the browser stores a local copy of it, allowing user edits without incurring network latency. The propagation of changes to all clients is managed by a variant of Operational Transformation (OT) [10, 29], the most common conflict-resolution approach for RTCE editors, whose details are not relevant for our discussion.

What is relevant is that each pad *revision* is described by metadata about the revision author and timestamp, as well as a machine-readable description of the *changeset* that led to that revision. A changeset is a sequence of low-level *edit operations* (or simply “edits”). In order to apply a changeset, Etherpad executes each of its edits in turn, modifying the current pad textual content and attributes as it goes.

The basic idea of our model is to classify individual edits as collaborative or not in time (RQ1), space (RQ2), or both (RQ3). Each edit can be independently classified as collaborative or not in each dimension. The overall collaboration on a given pad can then be calculated by simply counting the number of edits flagged as collaborative along each dimension.

Note that our model does not consider as collaborative those edits that are too distant — either in space or in time or both — even if they could be if we followed a broader definition of collaboration. The co-authorship itself, regardless of when and where the content is edited, is a form of collaboration. Our focus is not on that aspect. Following the 2x2 structure of the Johansen’s time-space matrix [18], we could have considered remote and non-collocated edits as a fourth pattern of collaboration (no space and no time proximity, but still co-authorship); we decided not to, in order to stress on the actual joint editing of content.

3.1 Time collaboration

For *time collaboration*, the idea is to consider an edit collaborative if it is “close enough in time to an edit applied by a *different* author”. The only edit-related information required to evaluate this property are the timestamp and author of each edit; the actual changes applied to the pad are irrelevant, and so are their spatial positions in the document.

Additionally, we need a global metric—which we call the *time window*—to capture the notion of being “close enough” (in time): an edit is in fact marked as collaborative if there exists at least one other edit that (a) was performed by a *different* author and (b) occurred within the chosen time window.

3.2 Space collaboration

Space collaboration measures the extent to which authors “worked on the same area of a pad”.

The idea is to maintain a running mapping from characters in the pad to the author who last modified it. This is similar to the coloring that Etherpad user interface uses to associate characters to authors. We replay the edit history and, whenever an edit is applied, we compare its author to the authors of impacted and nearby characters: if there is at least one character associated to a different author the edit is considered to be space-collaborative. To capture the notion of being spatially near we use a *space window* expressed as a number of characters w_s . The window is split in two equal parts which are added before and after the interval of the edit under consideration.

Note that intervals and edits are processed in different ways when re-building the pads histories depending on their type (for instance, insertions vs. deletions): details of such processing are provided in Appendix A.

Table 1. Experimental data: amount of extracted pads, changesets, and edits

pads (documents)	13 871
changesets	4 274 692
edits	6 065 234

3.3 Spacetime collaboration

Naturally spacetime collaboration combines the ideas of space- and time-collaboration in order to identify edits which happen both (almost) synchronously and on nearby pad areas. Note that this cannot be a simple logical AND conjunction of being space- and time-collaborative.

The correct way to model this is hence to mark an edit e as *spacetime collaborative* when it is both space- and time-collaborative with respect to the *same* edit e' by a different author.

4 EXPERIMENTAL METHODOLOGY

Using the collaboration model sketched in the previous section we can now *measure* the overall amount of time-, space-, and spacetime collaboration witnessed by any given pad. To do so it is enough to calculate the ratio of time-collaborative edits, space-collaborative, and spacetime-collaborative ones over the total number of pad edits.

We wrote a straightforward implementation of this method, available at <https://github.com/gdazero/RTCE/>, and run it on a large corpus of pads taken from the Etherpad instance operated by the Wikimedia foundation at <https://etherpad.wikimedia.org>, which is one of the most popular publicly-accessible Etherpad instances.² This choice of instance is significant not only due to its large number of pads and users, but also because it consists entirely of real-life editing sessions—between Wikimedia employees and random users—as opposed to controlled in-lab experiments like those discussed in Section 2.

The Wikimedia Technical Operations team provided us with an anonymized Etherpad database dump as of May 2015, in which authors and pad contents could not be recognized, but where the lengths of individual edit operations have been preserved. Table 1 summarizes the data we have extract from the initial dump and subsequently studied: more than 4 million changesets coming from about 14 thousand pads, whose edit histories were composed of 6 million individual edits. In comparison with the largest quantitative studies found in the literature, this dataset is two orders of magnitude larger in terms of the amount of analyzed documents.

Time windows. As the model is parametric in the size of space- and time-windows, we repeated the analysis varying the sizes of either of them. For the time window we used 5, 10, and 60 second intervals. Such relatively small values are functional to the goal of our analysis and, as such, much smaller than time intervals used in the past in the literature to study different aspects of collaboration.

Consider the goal of studying the duration of working sessions authors spent writing together. A common approach to address this is to mine periods of (in)activity from edit logs and aggregate edits into clusters, called *sessions* [13]. This approach was initially used to track browsing activities on the Web [22] and subsequently applied to the editing process, for instance on Wikipedia. The typical length of collaborative sessions is 30 minutes, with researchers finding larger time windows up to one hour [16]. These values are compatible with real world activities such taking notes during a meeting, or sprint sessions to finalize a paper. Joint editing sessions have also been studied using shorter time windows. Olson et al. [24] combined edits into sessions of 7 minutes and classified

²see the census of Etherpad instances at <https://github.com/ether/etherpad-lite/wiki/Sites-that-run-Etherpad-Lite>

Table 2. Statistical overview of pads in the dataset (13 871 pads)

	avg	median	max	stddev	IQR
Length (lifetime average)	875	443	50 211	1964	697
Edits	437	37	184 372	2344	373
Authors	1.91	1	79	2.81	1

them as “simultaneous” if two or more authors edited the document in that time interval. Sun et al.[30] considered 15 minutes as average length for collaborative working sessions.

Our goal is not to capture those aspects of time collaboration, but rather to study to what extent authors work (almost) *simultaneously* by exploiting RTCE capabilities. Thus, we selected relative small time windows. The rationale is that an interval lower than 5 seconds is too quick to be meaningful when considering typical network latency between Etherpad clients and servers; 5–10 seconds seem reasonable thresholds to consider two or more edits to be (almost) simultaneous; 1 minute is definitely too long by any reasonable definition of “simultaneous” but, when compared with the results obtained for the former values, it allows to study how time collaboration evolves with increasingly large time windows.

Space windows. For the space window we considered four intervals: 10, 80, 400, and 800 characters. The smallest value was selected to capture edits that happen on a single word or nearby words, assuming that most pads were written in English, whose average word length is 5 letters [21]. 80 characters correspond to the typical size of a full line of text (as it is the case for hard-wrapped lines in terminals) and was used as mid point in our range of intervals; 400- and 800-character windows were selected as representative of changes happening in the same block of 5 or 10 lines of texts, corresponding to medium to long paragraphs. These values are in line with guidelines for writing plain English, that suggest that the average length of a paragraph should be 100-150 words for ordinary writing [19] and decrease for online writing down to 50 words [7].

The values chosen for space windows are relatively small too, in order to extract data about edit locality, rather than data about the distribution of work on large chunks of text or entire sections. The latter has been studied by Olson et al., whose experiments showed that teams tend to assign entire sections to specific team members, so as to better coordinate team work.

In total, we obtain 12 experimental configurations: 3 time windows \times 4 space windows. For each of them we run the tool implementing the model to classify individual edits as collaborative or not, along each of the three dimensions of RQ1 to RQ3.

5 RESULTS

In this section we present the results of analyzing the pad corpus with the proposed collaboration model and experimental methodology.

Table 2 presents a statistical overview of pads in the dataset, reporting average, median, maximum, standard deviation and InterQuartile Range (IQR) for three measures: length (as average number of characters in the pad across all its edits), number of edits, and number of authors. Some edit histories contain edits with empty author, which are used by Etherpad to synchronize data among clients. In order to properly rebuild edit histories these edits cannot be entirely discarded, but in Table 2 only real edits (and real authors) are shown. Overall, there is significant heterogeneity in terms of pad length, edits, and authors. Arguably this is exactly what one would expect from a public, popular deployment of a RTCE platform.

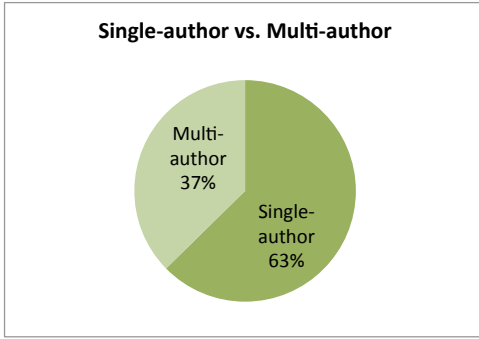


Fig. 1. Single-authored vs multi-authored pads

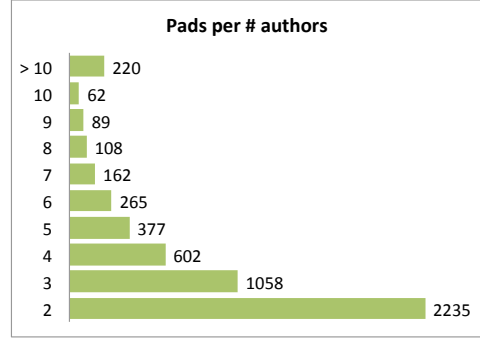


Fig. 2. Pads with a given number of authors

5.1 Multi-authorship

The first aspect we studied was the number of authors who edited each pad in the corpus.

Figure 1 shows the number of single-author vs multi-author pads. Authors tend to use Etherpad for personal editing activities. 63% of the pads (8690 in total) were edited by a single author, ruling out collaboration with other authors—at least as collaboration is observable from the RTCE point of view. (Note that according to our model having more than one author is a necessary but not sufficient condition: multiple authors might have worked on the same pad without having done any space, time, or spacetime collaborative edit. We will dig deeper into this aspect later on in this section.)

Figure 2 shows the distribution of pads with a given number of authors, focusing on pads edited by at least 2 authors (5181 pads). Within that subset, 2235 pads (43%) were edited by exactly 2 authors. The percentage grows to 82% (4273 pads) if we consider pads edited by at most 5 authors. It appears that, even when multiple authors work on a pad, collaboration groups remain relatively small. The amount of pads edited by larger groups indeed shrinks rapidly: only 686 pads (13%) were edited by 6–10 authors, only 220 (4%) by author groups larger than 10. Among those, three pads had more than 50 authors: one had 53, one 73, and one 79 (i.e., the maximum in the corpus).

When focusing on multi-author pads (see Table 3) we notice that heterogeneity persists, with standard deviations higher than the corresponding figures for the entire corpus (Table 2). Out of curiosity note that the longest pad has been edited by a single author.

5.2 Space collaboration

We now turn our attention to actual collaboration as captured by the model of Section 3. Table 4 summarizes our findings about space collaboration. Average, maximum, and standard deviation of

Table 3. Statistical overview of pads in the dataset, focusing on pads with at least 2 authors (5181 pads)

	avg	median	max	stddev	IQR
Length (lifetime average)	1472.15	794	44 020	2282.30	1490
Edits	1017.57	484	184 372	3716.28	943
Authors	3.99	3	79	3.71	2

Table 4. Amount of space collaboration

space win. / results	avg	max	stdev
10 characters	37.5 %	99.0 %	26.4 %
80 characters	49.6 %	99.9 %	30.1 %
400 characters	59.1 %	99.9 %	32.0 %
800 characters	61.7 %	99.9 %	32.2 %

the percentage of space-collaborative edits over all edits of a given pad are given, for various space window sizes.

In general the amount of space collaboration is high: from 37% for the 10 character window, up to 62% for 800 characters. Standard deviation values denote high heterogeneity. This is also in line with the maximum values of space-collaboration (minimum, not shown, being 0%), and the distribution of the number of edits. (We recall that these results should be read in light of the fact that we are considering only multi-author pads. If taken on the entire corpus, space collaboration averages would drop to 14–23%.)

The way space collaboration is measured is coherent with those relatively high percentages. Since an edit is marked as collaborative if it is “close enough” to any character that has been previously edited by another author, even minor changes affect the collaboration significantly. For example, the common action of reading what has been written by other authors and fixing the punctuation (or typos) generates a sequence of space collaborative edits.

Also, if an author adds a few characters at the end of the document and another author adds content before those characters, all these insertions are considered collaborative. This has a subtle consequence in presence of copy&paste operations. The paste of a large chunk of text is recorded by Etherpad as a sequence of contiguous edits of smaller size. If the content is pasted immediately before characters by a different authors, all those edits will be marked as collaborative. This is consistent with what want to capture, since the paste operation is close in space to content by others. This explains the maximum values observed for space collaboration (99%). In all those cases, which we investigated manually, authors pasted a long chunk of text immediately before content by a different author. Note that the values stabilize for a window of 80 chars: these pads are quite short and larger windows cross the boundaries, therefore cannot include further edits.

5.3 Time collaboration

The measured amount of time collaboration is reported in Table 5. Contrary to space collaboration, time collaboration average values are quite low, e.g., 8.5% with a 5 seconds time window. This is a

Table 5. Amount of time collaboration

time win. / results	avg	max	stdev
5s	8.5 %	98.2 %	15.7 %
10s	11.3 %	98.3 %	18.7 %
60s	21.9 %	99.6 %	27.4 %

Table 6. Amount of spacetime collaboration

time win. / space win.	10 chars	80 chars	400 chars	800 chars
5s	avg: 2.72 % max: 97.88 % stdev: 7.54 %	avg: 4.73 % max: 98.06 % stdev: 10.74 %	avg: 6.42 % max: 98.15 % stdev: 12.93 %	avg: 6.98 % max: 98.21 % stdev: 13.62 %
10s	avg: 3.90 % max: 97.88 % stdev: 9.39 %	avg: 6.62 % max: 98.09 % stdev: 13.43 %	avg: 8.67 % max: 98.23 % stdev: 15.73 %	avg: 9.35 % max: 98.3 % stdev: 16.44 %
60s	avg: 8.44 % max: 97.93 % stdev: 14.84 %	avg: 13.95 % max: 99.39 % stdev: 21.38 %	avg: 17.58 % max: 99.41 % stdev: 24.23 %	avg: 18.70 % max: 99.41 % stdev: 24.95 %

relevant finding: authors tend to collaborate much more in space than in time. Apparently, RTCE tools are used primarily for sharing documents and collaborate on them asynchronously.

The amount of collaborative edits increases with the window size, as expected, but it is worth noting that there is not a big change in collaboration when the time window is increased from 5 to 10 seconds (even though it is being doubled) while there is when its increased to 60.

The maximum values are also interesting. 98% of time-collaborative edits was unexpected. Digging deeper, we discovered that pads characterized by very high time collaboration also exhibit high space and spacetime collaboration. The number of authors editing these documents was very high too: one of them was modified by 8 authors, another by 14, with the most edited one counting up to 23 authors. This is a glaring, albeit infrequent, example of using RTCE tools for what they were designed for.

5.4 Spacetime collaboration

The results for spacetime collaboration are given in Table 6, for different pairs of space/time window sizes. Windows of 10 characters and 5 seconds denote real-time collaboration on nearby parts of the document by different authors; windows of 800 characters and 60 seconds denote loosely coupled collaboration, in both space and time. Results show that only a small part of the editing operations are spacetime collaborative (e.g., 2.7% for windows of 10 characters and 5 seconds). If we consider mid-size windows of 80 characters and 10 seconds, 6.6% of the editing operations are collaborative. This goes up to only 18.7% even with window sizes of 800 characters and 60 seconds.

The conclusion looks clear: the collaboration potential offered by RTCE tools is under-exploited in the studied corpus, apart very few notable examples.

6 DISCUSSION

In this section we review our findings in light of previous results in the literature and discuss potential implications for further research on RTCE.

6.1 Comparison with previous results

Under-exploitation of RTCE functionalities was previously observed by other researchers too. Chu and Kenney [27] compared wikis (MediaWiki) to RTCE tools (Google Docs) using user interviews. Interviewed authors were more willing to use RTCE tools for sharing and editing documents asynchronously, rather than in real-time. 70% of the 166 interviewed students in Brodhal et al. [6]

were not fully satisfied by the editing capabilities of the tested RTCE platforms. Both results look consistent with our findings that show a low amount of average collaboration.

The explanations given by unsatisfied testers of the second experiment are also consistent with our analysis. They claimed that it was difficult to keep track of document evolution in presence of multiple people editing synchronously and multiple changes being highlighted. This might explain the low amount of spacetime collaboration we observed: the cognitive overhead becomes too high when multiple authors edit the same document synchronously and in the same area.

The findings on the size of groups of collaborators outlined by Brodhal et al. are also comparable to ours. Most interviewed users suggested that teams of at most 2–3 authors are more productive than larger teams. The distribution of the number of authors we observed (Figure 2) is consistent with this: 63% of multi-author pads were edited by two (43%) or three (20%) authors. Larger teams seem to struggle in using RTCE technology effectively.

The experiment by Olson et al. [24] showed instead a high degree of collaboration, even using RTCE tools. Out of 96 documents edited by students over a 3-year period only 5 documents were edited by a single author, while in our findings most of the pads were single-authored. We believe that this is due to the highly controlled setting of Olson et al. experiment: students were asked to produce a team homework; as such they had to perform intrinsically collaborative tasks.

Students teams ended up being cohesive and well-coordinated: 77% of the documents were edited by all team members; the remaining ones only missed one or two members. The same analysis cannot be performed on our corpus since we have no knowledge of “team” structures. On the other hand, we can figure out which percentage of active authors was involved in at least one edit marked as collaborative. The findings confirm a lower degree of collaboration w.r.t. Olson et al. findings. With the smallest space window (10 chars) in 56,9% of the pads all authors collaborated in space (each author made at least one space-collaborative edit); the value raises to 61,6% with a space window of 800 characters, but it is still lower than Olson results. The same applies to time and spacetime collaboration. This is again explained by the controlled test setting. We observe that even in Olsen context space collaboration was higher than time collaboration: people mostly worked simultaneously on different parts of the document, with a divide-and-conquer strategy.

Another interesting analogy is that Olson et al. identified major bursts of activity and increased collaboration on deadline days. This could only be speculated on our dataset, but some peaks in the number of authors and collaborative edits make us think that some pads were indeed edited under time constraints and, due to that, they witnessed high collaboration.

Our data on the type and amount of edits are instead in contrast to the findings of Birnholtz et al. [5]. Their experiments showed that people make significant changes themselves and use communication tools (primarily comments and chat) to suggest small changes instead of making them directly. Our experiments showed a different behavior: a lot of small (collaborative) changes were found (making us suppose that Etherpad authors prefer to modify content directly instead of commenting or using other extra-document communication channels. We speculate this depends on the interface of the tool: Google Docs provides more usable and sophisticated tools for commenting and chatting while editing, compared to Etherpad.

6.2 Implications for further research

What are the implications of the findings presented in this paper for future work on RTCE collaboration and technologies?

First, the limited exploitation of collaboration as defined by our model raises questions about the most appropriate granularity of RTCE protocols. Is it really worth to use full-fledged Operational Transformation (OT) at character granularity, when splitting the documents in larger, individually lockable parts, would not get in the way of authors practices? Considering that most of the

collaboration happens in either space or time, but not in spacetime, our findings suggest that section-based conflict avoidance, as implemented for instance by MediaWiki, would be enough to cater for author needs. Intermediate synchronization granularities have been partially explored in the area of cryptographically secure pads [8, 34], but there are no final answers yet. Our experiments show that this design space for RTCE should be explored further to find a sweet spot between complexity and adhering to author needs.

Second, we observe that for the majority of our corpus (single-author pads), real-time content synchronization was not needed at all. Those documents could have even been edited while offline, with periodic one-way synchronization to the server (e.g., for backup assurance) when network was available. While offline Web application technologies are not yet widely standardized and supported by all browsers, their future increased availability might prove to be a very good fit for a wide range of editing scenarios that are currently being addressed by RTCE.

Third, the user experience of RTCE tools needs to be further improved. As discussed by others and confirmed by our results, there is a threshold effect in the number of authors that eventually collaborate. Solutions are needed for making changes clear to the authors, even when there are a lot to be presented, for instance by filtering changes by type or aggregating small changes into larger ones. The same suggestion was made by Birnholtz et al. [5], who claimed that *“features for explaining and presenting edits can focus on minor edits — [...] designers may wish to consider tools to support easier description and explanation of more major edits and restructuring”*. New user interface optimized for space collaboration could also be designed: they would allow users to work simultaneously in different parts of the document, minimizing visual distractions. Today users are provided with split or side-by-side windows but these solutions appear to be sub-optimal.

Fourth, our results give insights on how to design end-to-end encrypted RTCE tools, where users need not to trust platform hosters for content confidentiality [8, 34]. A key decision in designing such tools is whether block- or stream-ciphers should be used. Preliminary work by two authors of this paper hints at the fact that stream-ciphers would support collaboration at the cost of a deep redesign of conflict-resolution protocols and a significant implementation effort. Conversely, symmetric block-ciphers would be easy to implement but impose hard choices on the size of independently encrypted blocks and modes of operation by the chosen encryption algorithms. The network overhead introduced by a bad block size choice in presence of a significant amounts of spacetime collaboration might be very relevant. Even though more research is still needed, this paper findings on the (low) amount of collaboration in real-life RTCE support the block-cipher approach and provide useful figures to determine block sizes.

Finally, our analysis showed that there is need to study RTCE on rich(er)-text documents. Let's consider an HTML table. A naive application of the model described in this paper would consider the deletion of a whole column as a sequence of small edits spatially scattered throughout the document. It would be helpful to also investigate spatial metrics that allow to capture locality when working on tables, records, and structured data.

7 THREATS TO VALIDITY

We discuss in this section the potential threats to the validity of our experiment and findings, starting with internal validity aspects. The document corpus we analyzed has been provided by Wikimedia operations team (TO). We have no first-hand guarantee that the document histories stored in there actually correspond to all user edit actions. On the other hand, we expect histories to be correct as they are successfully used in production by Wikimedia Etherpad installation to present edit histories to their users.

Etherpads users are identified by their navigation sessions, stored using browser cookies. As a consequence, users will no longer be recognized as previous authors if they remove the relevant

cookies from their browser, or open the same pad from different machines or browsers. As a consequence our analysis might have characterized as collaborative edits performed by the same author. There is no way to verify this effect as we do not have access to information that would identify the original navigation sessions. On the other hand switching between browsers (or machines) is not very frequent.

Regarding space collaboration, we verify for each edit whether or not a given interval in the pad contains at least one character *most recently* modified by a different author. One could alternatively keep track of *all authors* who have modified any given character, with an approach similar to that used for measuring the author's contributions in wikis [1, 9]. Doing so would likely inflate the number of edits marked as space and spacetime-collaborative. Since this discrepancy only affects characters which are modified many times by different authors, we expect this to not significantly impact our findings. It is however an interesting direction that we plan to investigate in the future.

The proposed model also does not take into account deleted content when assessing space collaboration. This might again underestimate collaboration. However, deleted content is often followed by new characters inserted by the same author, because authors tend to *revise* content—by adding and removing snippets—instead of just removing characters. Given that inserted characters are taken into account when measuring space collaboration, this potential issue gets mitigated.

Regarding external validity we start by observing that our selections of space- and time-window values is debatable. We motivated our choices on the basis the average lengths of textual features such as words, phrases, and paragraphs. There is no evidence that those values are the best choices to characterize collaboration. However, as we performed a parameter sweep using multiple values for each window showing how collaboration is affected, we consider the results insightful nonetheless. Besides, the proposed *model* for quantitatively measure collaboration is independent from the actual window sizes: the proposed experiments can easily be repeated, possibly in different contexts, varying window sizes at will.

Furthermore, we do not claim that the findings presented here would necessarily apply *ipso facto* to RTCE platforms other than Etherpad. We expect the model to be general enough to be applied to most RTCE textual editors, and in particular to popular technologies such as Google Docs. We also expect the relatively low collaboration ratio to be common to settings similar to ours (i.e., freely accessible documents) no matter the chosen RTCE technology, but we cannot rule out the impact of extra-content communication mechanisms (e.g., comments, chat messages, etc.) we plan to explore these aspects as future work.

8 CONCLUSIONS

We have studied documents edited on one of the largest public instances (14 000 pads) of the Etherpad real-time collaborative editor (RTCE) to characterize real-time collaboration on the Web. The characterization is based on a novel model that captures collaboration along three dimensions: time (whether collaboration happens simultaneously or almost so), space (whether collaboration happens in nearby parts of the document), and spacetime (both).

Our findings show real-time collaboration happens, but not very often. A very limited number of pads witness simultaneous collaboration on nearby document areas; time collaboration (on unrelated document parts) is more frequent; space collaboration on the same document portions (but during editing sessions that are far apart in time) is the most common. These results are influenced by the public setting of the studied Etherpad instance. It is possible that private deployments would show larger amounts of collaboration within more coordinated teams working on shared documents. On the other hand the study of a freely accessible service shed some light onto real-time collaboration in the wild.

ACKNOWLEDGMENTS

We would like to thank the Technical Operations team at the Wikimedia Foundation for their help and continued support in retrieving the anonymized document corpus without which this work would not have been possible.

A CHARACTERIZING REAL-TIME COLLABORATION ON ETHERPAD

This section describes in detail the conceptual model that we used to classify individual edits as collaborative or not in time (RQ1), space (RQ2), or both (RQ3). To do so, we first provide some more details on the internals of Etherpad to the extent needed to understand the proposed approach.

A.1 Etherpad and basic notation

Etherpad [12] is a web-based real time collaborative editor implemented using the Node.js framework. It supports textual documents only, called *pads*, which users can collaboratively edit directly from within web browsers. Semantically a pad is a long textual string, possibly containing newline separators to separate blocks, together with a set of attributes (e.g., bold, italic, etc.) that apply to selected sub-strings of the pad textual content.

When a pad is opened the browser stores a local copy of it, allowing user edits without incurring network latency. Local changes are propagated in real-time to all other connected clients, routing through a centralized server, the same server that hosts the Etherpad web application.

To avoid conflicts and maintain editing consistency, Etherpad implements its own variant of Operational Transformation (OT) [10, 29], which is the most common conflict-resolution approach for RTCE editors.

The details of Etherpad's own OT variant are not relevant for our discussion. What is relevant is that each pad *revision* is captured as a JSON object within Etherpad, containing metadata about the revision author and timestamp (in the *meta* field), as well as a machine-readable description of the *changeset* that led to that revision.

A changeset is a sequence of low-level *edit operations* (or simply "edits"). In order to apply a changeset, Etherpad executes each of its edits in turn, modifying the current pad textual content and attributes as it goes. Four types of edits are supported by Etherpad:

Ins(C,A) insert string *C* with text attributes *A* at the current position

Del(l) delete *l* characters, starting from the current position

Cpy(l) preserve *l* characters (in both their textual value and attributes), starting from the current position

Upd(l,A) update (i.e., set) the attributes of the next *l* characters to match attributes *A*

The application of a changeset starts at the beginning of the pad. Each edit moves the current position forward, with the exception of **Del** which deletes characters in front of the cursor without changing its position. The **Cpy** operation can therefore be used to "move the cursor" and perform a subsequent edit operation at the desired position.

An important difference between **Cpy** and **Upd** is that while **Upd** corresponds to explicit user actions (e.g., typeset the word "Example" in italic), **Cpy** does not, it is an Etherpad artifact to avoid having to specify absolute positions for every edit. Note also that each edit is applied on the pad as modified by the previous edits. Let's consider the edits: **Del**(226), **Ins**("ab"). The first deletes 226 characters starting from position 1; the second one inserts the string "ab" starting again from position 1, since **Del** just removed some characters without moving the cursor.

Let us now go into the details of our model. Throughout the rest of the section, we will refer to the edit history shown in Figure 3.

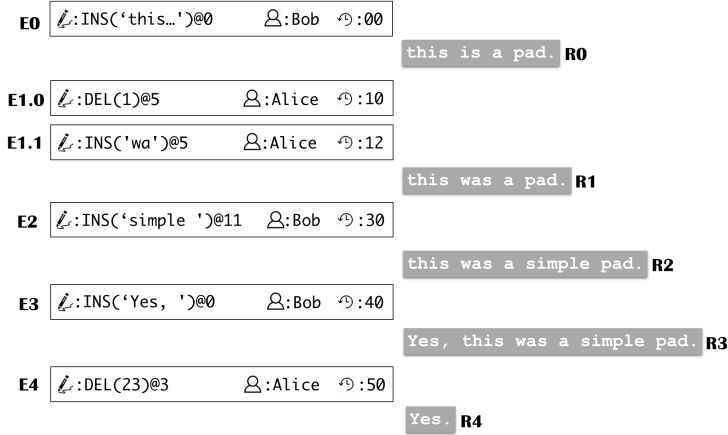


Fig. 3. Sample Etherpad edit history comprising 5 revisions generated from 6 edits by 2 authors

The pad contains 5 revisions (numbered R0–R4) generated from 5 changesets (E0–E4); one of them is composed of two edits (E1.0 and E1.1), the others of singleton edits. For each changeset the figure shows: the name of the author (Alice or Bob), the timestamp (where 0 indicates document creation time) and the edits it comprises. The right hand side of the picture shows the full pad content after each edit. For the sake of brevity we omit Cpy edits (which are no-op) and show the position of each edit after @.

The following notation elements will be used throughout the section:

interval $p[start, end]$ indicates the sequence of characters in the pad p , starting at position $start$ (0-based) and ending at position end . E.g., in pad R0 $p[3, 13]$ is the string "s is a pad" and all $p[n, n]$ intervals are empty strings.

length $length(p)$ is the number of characters in pad p .

timestamp $time(e)$ is the timestamp of edit e .

author $author(e)$ is the string identifying the author of edit e .

A.2 Time collaboration

Each edit can be independently classified as collaborative or not in space, time, and spacetime (which, as it will be detailed in Section 3.3, is more subtle than being *both* space- and time-collaborative).

For *time collaboration*, the idea is to consider an edit collaborative if it is “close enough in time to an edit applied by a *different* author”. The only edit-related information required to evaluate this property are the timestamp and author of each edit; the actual changes applied to the pad are irrelevant, and so are their spatial positions in the document. Additionally, we need a global metric—which we call the *time window*—to capture the notion of being “close enough” (in time).

An edit is marked as collaborative if there exists at least one other edit that (a) was performed by a *different* author (i.e., you cannot “collaborate with yourself”) and (b) occurred within the chosen time window. Formally:

DEFINITION 1 (TIME COLLABORATION). *An edit e_1 is collaborative in time (or time-collaborative) with respect to time window w_t iff*

$$\exists e_2 \text{ s.t. } time(e_2) \in [time(e_1) - w_t, time(e_1)] \wedge author(e_1) \neq author(e_2)$$

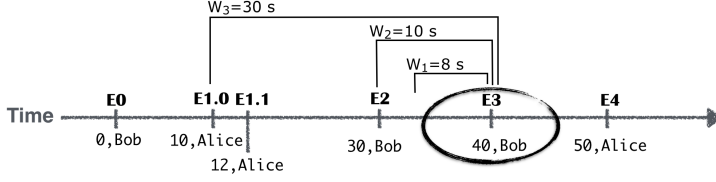


Fig. 4. Timeline of Figure 3 edits. Edit E3 is time-collaborative when using time window W_3 (30 seconds), but is not with windows W_1 (8s) and W_2 (10s).

Table 7. Definitions of pre- and post-intervals for all edit operations

edit @ position	pre-interval	post-interval
Ins(C, A)@ p	$[p, p]$	$[p, p + C]$
Del(l)@ p	$[p, p + l]$	$[p, p]$
Upd(l, A)@ p	$[p, p + l]$	$[p, p + l]$
Cpy(l)@ p	n/a	n/a

Consider Figure 4, which provides a time-indexed visualization of the edits in Figure 3. Edit E3 it is not time-collaborative w.r.t. time window $w_1 = 8$ seconds since there is no edit looking behind 8 seconds; not it is collaborative w.r.t. $w_2 = 10$ seconds either, since the only edit in the time window is by the same author. It is however collaborative w.r.t. $w_t = 30$ seconds, since there are two edits in that time window by author Alice.

Note that the model relies on an implicit relationship between pairs of edits: each collaborative edit is connected to one (or more) other edit(s) that caused it to be marked as collaborative. Such a relation is not symmetric: the fact that e_1 is collaborative because of e_2 does not imply that e_2 is collaborative. We wanted to be able to measure collaboration at any point in time of pad evolution; marking an edit as collaborative because of a *subsequent* would have made that impossible. A corollary of this is that the first edit of a pad cannot be collaborative.

A.3 Space collaboration

Space collaboration measures the extent to which authors “worked on the same area of a pad” (the *where*). To model this we first need to define what is an “area” of a pad and which are its boundaries. The basic idea is to identify the characters which are affected by a given edit, distinguishing between the characters on which the edit “started” and those “generated” by it, as captured by the following notions:

pre-interval: the interval denoting the characters which are in the pad when a given edit is performed and that will be affected by that edit.

post-interval: the interval denoting the characters affected by an edit and that remain present in the pad after the application of that edit.

Pre- and post-intervals depend on the nature of the edit. Ins edits have empty pre-intervals (they do not affect existing characters) and non-empty post-intervals (the inserted characters); conversely Del edits have non-empty pre-intervals (the characters to be deleted) and empty post-intervals. Table 7 gives the precise definitions of pre- and post-intervals for all pad edit operations. Note that Upd edits have the same values for both intervals since they do not add to or remove characters from the pad. Note also that intervals are not meaningful for Cpy, since that operation does not alter the pad.

Pre- and post-intervals are the building blocks used to model space collaboration. An edit is *space collaborative* if it is “close enough” (in space) to an edit of a *different* author.

To assess whether an edit is space-collaborative or not we maintain a running mapping from characters in the pad to the author who last modified it. We replay the edit history and, whenever an edit is applied, we compare its author to the authors of impacted and nearby characters: if there is at least one character associated to a different author the edit is considered to be space-collaborative. To capture the notion of being spatially near we use a *space window* expressed as a number of characters w_s . The window is split in two equal parts which are added before and after the pre-interval of the edit under consideration.

Consider the example in Figure 5. White background characters are associated to Bob; gray background ones to Alice. Edit E2 consists of an Ins by Bob at position 11. Using a space window $W_1 = 6$ characters, the pre-interval $[11, 11]$ is expanded to $[8, 14]$. Since all characters in the resulting interval belong to Bob there is no space collaboration. Using a space window $W_2 = 10$ characters on the other hand results in marking E2 as collaborative: the pre-interval $[11, 11]$ is expanded to $[6, 16]$, which includes character “a” by Alice.

Formally, we define:

lastedit at a given time t , $lastedit(pad, p)$ denotes the most recent edit whose post-interval includes the character of pad at position p

expand $expand(pad[l, r], w) = pad[max(0, l - w), min(r + w, length(pad))]$

using which we can formalize the notion of space collaboration as:

DEFINITION 2 (SPACE COLLABORATION). An edit e_1 is collaborative in space (or *space collaborative*) with respect to space window w_s iff

$$\exists c \text{ s.t. } c \in expand(pre-interval(e_1), w_s) \wedge author(lastedit(c)) \neq author(e_1)$$

i.e., an edit is space-collaborative when there is at least one character in its pre-interval, expanded using the space window, that has been most recently modified by a different author.

The overall space collaboration on a given pad (RQ2) can then be calculated by simply counting the number of edits flagged as space collaborative according to the above definition.

A.4 Spacetime collaboration

A formalization of the notion of *spacetime* collaboration completes the model.

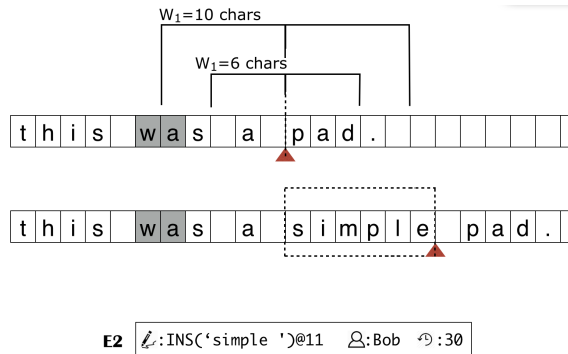


Fig. 5. Space collaboration for edit E2 on the pad of Figure 3. The edit is not flagged as space collaborative using a space window $W_1 = 6$ characters, but it is using a space window $W_2 = 10$ characters.

Naturally spacetime collaboration combines the ideas of space- and time-collaboration in order to identify edits which happen both (almost) synchronously and on nearby pad areas. The formal definition is hence derived from previous definitions, but is more subtle than a simple logical AND conjunction of being space- and time-collaborative.

Consider an edit e_1 that is marked as time-collaborative due to a preceding edit e_2 which falls within e_1 's time window, but has happened on a pad area far apart from the pre-interval of e_1 , escaping the chosen space window. The same edit e_1 is also marked as space-collaborative, but due to *another* edit $e_3 \neq e_2$: e_1 and e_3 affect nearby areas of the pad, but happened far apart in time. In this case e_1 should not be considered spacetime-collaborative, because there is no single edit by another author that happened both at around the same time and in a nearby document area.

The correct way to model this is hence to mark an edit e as *spacetime collaborative* when it is both space- and time-collaborative with respect to the *same* edit e' by a different author. This idea is formally captured in the concluding definition of our model of collaboration:

DEFINITION 3 (SPACETIME COLLABORATION). *The edit e_1 is collaborative in spacetime (or spacetime-collaborative) with respect to time window w_t and space window w_s iff:*

$$\begin{aligned} \exists c \text{ s.t. } & c \in \text{expand}(\text{pre-interval}(e_1), w_s) \\ & \wedge \text{author}(\text{lastedit}(c)) \neq \text{author}(e_1) \\ & \wedge \text{time}(\text{lastedit}(c)) \in [\text{time}(e_1) - w_t, \text{time}(e_1)] \end{aligned}$$

REFERENCES

- [1] B. Thomas Adler, Luca de Alfaro, Ian Pye, and Vishwanath Raman. 2008. Measuring Author Contributions to the Wikipedia. In *Proceedings of the 4th International Symposium on Wikis (WikiSym '08)*. ACM, New York, NY, USA, Article 15, 10 pages. <https://doi.org/10.1145/1822258.1822279>
- [2] Mehdi Ahmed-Nacer, Claudia-Lavinia Ignat, Gérald Oster, Hyun-Gul Roh, and Pascal Urso. 2011. Evaluating CRDTs for real-time document editing. In *Proceedings of the 2011 ACM Symposium on Document Engineering*. ACM, 103–112. <https://doi.org/10.1145/2034691.2034717>
- [3] Paul André, Robert E. Kraut, and Aniket Kittur. 2014. Effects of Simultaneous and Sequential Work Structures on Distributed Collaborative Interdependent Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 139–148. <https://doi.org/10.1145/2556288.2557158>
- [4] Jeremy Birnholtz and Steven Ibara. 2012. Tracking Changes in Collaborative Writing: Edits, Visibility and Group Maintenance. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 809–818. <https://doi.org/10.1145/2145204.2145325>
- [5] Jeremy P. Birnholtz, Stephanie B. Steinhart, and Antonella Pavese. 2013. Write here, write now!: an experimental study of group maintenance in collaborative writing. In *2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '13*. ACM, 961–970. <https://doi.org/10.1145/2470654.2466123>
- [6] Cornelia Brodahl and Nils Kristian Hansen. 2014. Education Students' Use of Collaborative Writing Tools in Collectively Reflective Essay Papers. *Journal of Information Technology Education: Research* 13 (2014), 91–120. <http://www.jite.org/documents/Vol13/JITEv13ResearchP091-120Brodahl0463.pdf>
- [7] Martin Cutts. 2013. *Oxford guide to plain English*. OUP Oxford.
- [8] Gabriele D'Angelo, Fabio Vitali, and Stefano Zacchiroli. 2010. Content Cloaking: Preserving Privacy with Google Docs and Other Web Applications. In *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC '10)*. ACM, 826–830. <https://doi.org/10.1145/1774088.1774259>
- [9] Luca de Alfaro and Michael Shavlovsky. 2013. Attributing Authorship of Revisioned Content. In *Proceedings of the 22Nd International Conference on World Wide Web (WWW '13)*. ACM, New York, NY, USA, 343–354. <https://doi.org/10.1145/2488388.2488419>
- [10] C. A. Ellis and S. J. Gibbs. 1989. Concurrency Control in Groupware Systems. In *Proceedings of the 1989 ACM SIGMOD International Conference on Management of Data (SIGMOD '89)*. ACM, 399–407. <https://doi.org/10.1145/67544.66963>
- [11] J.H. Erik Andriessen. 2002. *Working with Groupware: Understanding and Evaluating Collaboration Technology*. Springer.
- [12] Etherpad Foundation. 2018. Etherpad. <http://etherpad.org/>
- [13] R. Stuart Geiger and Aaron Halfaker. 2013. Using Edit Sessions to Measure Participation in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 861–870. <https://doi.org/10.1145/2441776.2441873>

- [14] Google Inc. 2018. Google Docs. <http://docs.google.com/>
- [15] Jonathan Grudin. 1994. Computer-Supported Cooperative Work: History and Focus. *Computer* 27, 5 (May 1994), 19–26. <https://doi.org/10.1109/2.291294>
- [16] Aaron Halfaker, Oliver Keyes, Daniel Kluver, Jacob Thebault-Spieker, Tien Nguyen, Kenneth Shores, Anuradha Uduwage, and Morten Warncke-Wang. 2015. User Session Identification Based on Strong Regularities in Inter-activity Time. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 410–418. <https://doi.org/10.1145/2736277.2741117>
- [17] Claudia-Lavinia Ignat, Gérald Oster, Olivia Fox, Valerie L. Shalin, and François Charoy. 2015. How Do User Groups Cope with Delay in Real-Time Collaborative Note Taking. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*, Nina Boulus-Rødje, Gunnar Ellingsen, Tone Bratteteig, Margunn Aanestad, and Pernille Bjørn (Eds.). Springer International Publishing, 223–242. https://doi.org/10.1007/978-3-319-20499-4_12
- [18] Robert Johansen. 1988. *GroupWare: Computer Support for Business Teams*. The Free Press, New York, NY, USA.
- [19] Margaret H Larock, Jacob C Tressler, and Claude E Lewis. 1980. *Mastering Effective English*. Copp Clark Pitman.
- [20] Du Li and Rui Li. 2006. A Performance Study of Group Editing Algorithms. In *Proceedings of the 12th International Conference on Parallel and Distributed Systems - Volume 1 (ICPADS '06)*. IEEE Computer Society, Washington, DC, USA, 300–307. <https://doi.org/10.1109/ICPADS.2006.18>
- [21] Mark S Mayzner and Margaret Elizabeth Tresselt. 1965. Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic monograph supplements* (1965).
- [22] Alan L. Montgomery, Shibo Li, Kannan Srinivasan, and John C. Liechty. 2004. Modeling Online Browsing and Path Analysis Using Clickstream Data. *Marketing Science* 23, 4 (Sept. 2004), 579–595. <https://doi.org/10.1287/mksc.1040.0073>
- [23] Ricardo Olenewa, Gary M. Olson, Judith S. Olson, and Daniel M. Russell. 2017. Now That We Can Write Simultaneously, How Do We Use That to Our Advantage? *Commun. ACM* 60, 8 (July 2017), 36–43. <https://doi.org/10.1145/2983527>
- [24] Judith S. Olson, Dakuo Wang, Gary M. Olson, and Jingwen Zhang. 2017. How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Trans. Comput.-Hum. Interact.* 24, 1, Article 4 (March 2017), 40 pages. <https://doi.org/10.1145/3038919>
- [25] Gérald Oster, Pascal Urso, Pascal Molli, and Abdessamad Imine. 2006. Data Consistency for P2P Collaborative Editing. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 259–268. <https://doi.org/10.1145/1180875.1180916>
- [26] I. R. Posner and R. M. Baecker. 1992. How people write together [groupware]. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, Vol. iv. 127–138 vol.4. <https://doi.org/10.1109/HICSS.1992.183420>
- [27] David M. Kennedy Samuel Kai-Wai Chu. 2011. Using Online Collaborative Tools for Groups to Co-Construct Knowledge. *Online Information Review* 35, 4 (2011), 581–597. <https://doi.org/10.1108/14684521111161945>
- [28] Bin Shao, Du Li, Tun Lu, and Ning Gu. 2011. An Operational Transformation Based Synchronization Protocol for Web 2.0 Applications. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. ACM, 563–572. <https://doi.org/10.1145/1958824.1958910>
- [29] Chengzheng Sun and Clarence A. Ellis. 1998. Operational Transformation in Real-Time Group Editors: Issues, Algorithms, and Achievements. In *CSCW '98, Proceedings of the ACM 1998 Conference on Computer Supported Cooperative Work*. ACM, 59–68. <https://doi.org/10.1145/289444.289469>
- [30] Yunting Sun, Diane Lambert, Makoto Uchida, and Nicolas Remy. 2014. Collaboration in the Cloud at Google. In *Proceedings of the 2014 ACM Conference on Web Science (WebSci '14)*. ACM, New York, NY, USA, 239–240. <https://doi.org/10.1145/2615569.2615637>
- [31] David Wang, Alex Mah, and Soren Lassen. 2015. *Apache Wave (incubating) Protocol Documentation, Release 0.4*. Technical Report. The Apache Software Foundation. https://people.apache.org/~al/wave_docs/ApacheWaveProtocol-0.4.pdf
- [32] Dakuo Wang, Judith S. Olson, Jingwen Zhang, Trung Nguyen, and Gary M. Olson. 2015. DocuViz: Visualizing Collaborative Writing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. New York, NY, USA, 1865–1874. <https://doi.org/10.1145/2702123.2702517>
- [33] Dakuo Wang, Haodan Tan, and Tun Lu. 2017. Why Users Do Not Want to Write Together When They Are Writing Together: Users' Rationales for Today's Collaborative Writing Practices. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 107 (Dec. 2017), 18 pages. <https://doi.org/10.1145/3134742>
- [34] Chunwang Zhang, Junjie Jin, Ee-Chien Chang, and Sharad Mehrotra. 2012. Secure Quasi-Realtime Collaborative Editing over Low-Cost Storage Services. In *Secure Data Management - 9th VLDB Workshop, SDM 2012 (Lecture Notes in Computer Science)*, Vol. 7482. Springer, 111–129. https://doi.org/10.1007/978-3-642-32873-2_8

Received April 2018; revised July 2018; accepted September 2018